基于信号检测论的错误信息鉴别层级模型

曹呈旭　七十三　金童林　曾小叶　安叶青　卜塔娜

(内蒙古师范大学心理学院, 内蒙古高校心理教育重点研究基地, 呼和浩特 010022)

摘　要　在错误信息鉴别的研究领域，系统 2 动机性推理理论和经典推理理论分别从不同视角探讨了影响个体错误信息鉴别的因素，但两者在认知能力的作用解释上存在分歧。在现有研究基础之上，引入情绪、信息特征和个体立场及其深层次动机等因素，进一步完善基于信号检测论的错误信息鉴别层级模型，旨在深化对不同因素如何影响错误信息鉴别的理解。该模型通过区分不同因素对信息鉴别中辨别敏感性和判断标准的影响，不仅有效地调和了系统 2 动机性推理理论和经典推理理论在认知能力作用观点上的分歧，也为理解错误信息鉴别的复杂机制提供了更为细致和结构化的分析框架。

关键词　错误信息, 信号检测论, 影响因素, 系统 2 动机性推理理论, 经典推理理论

分类号　B849

1　引言

网络通讯技术发展带来诸多便利的同时，也滋生了一些危害，错误信息[1]便是其中之一(彭知辉, 2022; 吴诗苑 等, 2022; Wang et al., 2019; Wardle & Derakhshan, 2017)。与真实信息相比，错误信息的传播更快、更深、更远、更广(Vosoughi et al., 2018)。其不仅会影响个体对事件做出最优决策，还可能引发诸多社会负面效应(Lewandowsky et al., 2017)。因此，一些研究从不同视角探索了个体为何会相信错误信息(Brashier & Marsh, 2020; Pennycook & Rand, 2021)，并提出了如何鉴别和纠正错误信息的建议(Lewandowsky et al., 2012; Swire et al., 2017)。

信息鉴别是一个复杂的过程(Brashier & Marsh, 2020)，其会受到认知、情绪和动机等多

收稿日期: 2023-09-15
内蒙古师范大学研究生科研创新基金资助项目(CXJJB23002)。
通信作者: 七十三, E-mail: Qshisan@126.com

[1] 在阅读相关文献时，我们注意到诸如“虚假信息”、“错误信息”、“不良信息”、“假新闻”、“谣言”等术语常常被混用，造成了研究概念界定的模糊和研究内容的混淆。一些研究者指出，“错误信息”(misinformation)这一术语可分为狭义和广义两种解释。狭义的“错误信息”指无意传播的虚假的或者不准确的信息(彭知辉, 2022; Wardle & Derakhshan, 2017)。广义的“错误信息”作为一个总括性术语，包含了虚假信息在内的谣言、假新闻、阴谋论等所有可能导致严重后果的不实信息(吴诗苑 等, 2022; Wang et al., 2019)。鉴于此，本文采用“错误信息”的广义定义，以便更全面地覆盖和探讨相关研究。

种因素的影响。这些因素之间的复杂交互作用会导致不同理论对影响因素的解释存在差异。例如，系统 2 动机性推理理论和经典推理理论在解释个体认知能力对错误信息鉴别的作用方面存在不同观点(Baron, 2017; Gawronski et al., 2023; Kahan & Corbin, 2016; Pennycook & Rand, 2019)。这种理论上的分歧，源于未能充分区分和认识不同因素对错误信息鉴别的作用。作为心理学研究中广泛采用的分析框架，信号检测论为区分和深入理解影响个体判断的各种因素提供了有效方法(Green & Swets, 1966; Lynn & Barrett, 2014)。基于此，本文在现有研究基础上，试图补充和完善基于信号检测论的错误信息鉴别层级模型，进一步阐释认知、情绪、动机以及信息特征等因素如何影响错误信息鉴别，以期调和现有理论之间的分歧，为后续研究提供更为细致的分析框架，并对实际应用中的信息鉴别和治理提供理论支持。

2　系统 2 动机性推理理论与经典推理理论是对立还是误会？

2.1　系统 2 动机性推理理论

个体在信息鉴别时并非一张白纸。动机性推理理论认为，个体的先前信念、期望和目标等因素均会影响其判断与决策(Kunda, 1990)。Kahan 等人通过系列研究进一步提出了系统 2 动机性推理理论(motivated System 2 reasoning account, MS2R)。该理论认为，认知能力越高的个体在信息鉴别时更倾向于信任和支持与自身利益相一致的观点(Kahan, 2013;Kahan & Corbin, 2016; Kahan et al., 2012; Kahan et al., 2017)。

Kahan 等人(2012)探究了科学素养和数字计算能力对气候变化风险感知的影响。结果发现，个体的科学素养和数字计算能力与气候变化风险感知呈负相关。但随着科学素养和数字计算能力的提高，个体对气候变化风险感知呈两极分化，更偏向于信任和维护自己所属群体的观点。这表明，个体的科学素养和数字计算能力越高，他们在信息鉴别时越倾向于信任和维护自己所属群体的观点。此外，Kahan (2013)的研究发现，分析性思维高的个体在信息鉴别时也同样会倾向于信任和支持他们所属群体的观点。

Kahan 等人(2017)的研究通过两项不同的信息鉴别任务，再次探讨了个体的认知能力与党派偏见的关系。其中一项任务是评估护肤霜治疗皮疹的有效性，该任务与党派偏见无关。另一项任务是评估禁枪降低城市犯罪率的有效性，该任务与党派偏见有关。结果显示，高计算能力的个体在评估护肤霜效用任务中表现出更高的准确性。然而，在评估禁枪效用任务中，

高计算能力的个体却表现出更高的党派偏见，这再次证明了系统 2 动机性推理理论。

2.2 经典推理理论

决策研究者常常将思维分为两个算法系统或两个处理过程(Evans & Stanovich, 2013; Thompson & Johnson, 2014; Tversky & Kahneman, 1983)。其中系统 1 是依赖直觉进行的启发式快加工，而系统 2 是依赖理性的分析性慢加工。经典推理理论(classical reasoning account)认为，当个体使用系统 1 进行信息鉴别时，会忽略信息细节进行粗略的估计和判断，进而降低了信息鉴别的准确性(Pennycook & Rand, 2019, 2020)。相反，采用系统 2 进行分析性推理时，能够提高个体信息鉴别的准确性。相关研究从个体差异层面(Pennycook & Rand, 2019, 2020)以及个体内部实验操纵层面表明，采用系统 2 分析性推理有助于提高信息鉴别的准确性。

Pennycook 和 Rand (2019)的研究探讨了分析性推理对错误信息鉴别的影响。研究结果反驳了系统 2 动机性推理理论，而支持了经典推理理论，即无论个体的政治立场如何，分析性推理都有助于提高错误信息鉴别的准确性。随后，他们的研究再次发现，分析性推理与错误信息鉴别之间的关系相对稳定，不受信息来源和信息熟悉度的影响(Pennycook & Rand, 2020)。同样，Pehlivanoglu 等人(2022)的研究也表明，具有较高分析性推理能力的个体在鉴别与 COVID-19 无关的错误信息时更准确，而且这种关系不受年龄的影响。这些个体差异层面的研究结果一致表明，分析性推理倾向高的个体在信息鉴别时更准确(Pennycook & Rand, 2019, 2020)。

此外，还有一些研究探讨了个体内部认知加工深度对错误信息鉴别的影响。Qiu 等人(2017)的研究发现，低质量信息过载和个体注意不足都会降低信息鉴别的准确性。Pennycook 等人(2020)通过在信息中添加警告标签的方式，促使个体投入更多的分析性推理来鉴别信息。结果发现，这种方式不仅提高了个体信息鉴别的准确性，也有效降低了个体对党派立场信息的信任。该结果同样支持了经典推理理论，并反驳了系统 2 动机性推理理论。Bago 等人(2020)采用了两次反应范式来探究分析性思维对信息鉴别的影响。结果发现，当个体有更多时间进行分析性思考时(Bago & De Neys, 2017, 2019)，减少了对信息错误的信任。此外，他们的研究也再次证实，分析性思维较高的个体在信息鉴别时更准确。这些个体内部实验操纵层面的研究表明，无论是提供充足的时间促使个体进行分析性推理(Bago et al., 2020)，还是通过在

信息中添加警告标签以激发分析性推理(Pennycook et al., 2020)，均能提高个体信息鉴别的准确性。

2.3 对立还是误会？

现有研究中，系统 2 动机性推理理论和经典推理理论常被视为相互对立的观点(Bago et al., 2020; Gawronski et al., 2023; Pehlivanoglu et al., 2022; Pennycook & Rand, 2019, 2020)。两个理论的核心争议在于：系统 2 动机性推理理论主张认知能力高的个体在信息鉴别时会更倾向于信任与自己党派立场一致的信息；而经典推理理论则认为，无论信息是否与个体的党派立场一致，个体的分析性思维越高，信息鉴别的准确性就会越高(李艳红, 刘佳诺, 2022)。虽然一些研究将这两个理论视为相互对立，但这种观点仍需细致推敲。

一方面，相关理论研究中使用的测量工具与其理论所指的对应不够清晰，这可能是导致结论矛盾的重要原因之一。越来越多的理论和模型将智力因素细分为认知能力和思维倾向两个方面(冯雪, 彭凯平, 2015; Stanovich, 2016)，但与两个理论相关的现有研究似乎未能充分区分这两个方面。特别是一些研究中将与数字计算能力紧密相关的测量工具(Thomson & Oppenheimer, 2016)与分析性推理等同起来(Bago et al., 2020; Kahan, 2013; Kahan, 2017; Kahan & Corbin, 2016; Kahan et al., 2012; Kahan et al., 2017; Pennycook & Rand, 2019, 2020)，或笼统地将这些因素归为认知能力(Kahan & Corbin, 2016)，这会导致错误地理解这些因素的作用，并引发理论间的矛盾。

更重要的是，两个理论关注的重点是不同的。系统 2 动机性推理理论聚焦于个体的社会性动机对错误信息鉴别的影响；而经典推理理论则侧重于个体的认知能力对错误信息鉴别的影响。将社会性动机与认知能力作为竞争性因素进行讨论似乎并不恰当，因为两者在错误信息鉴别时发挥的作用是不同的。认知能力代表的是个体完成任务的基础能力，其本身并不具备明确的目标指向性；而动机作为激发个体行为的心理趋向或内在驱力，具有明确的目标指向性。因此，更为恰当的方式是比较不同类型的动机，或比较不同的认知能力，以更准确地理解这些因素对错误信息鉴别的影响。例如，Rathje 等人(2023)的研究便探讨了个体在信息鉴别时社会性动机与准确性动机的竞争关系。研究发现，通过金钱激励手段对准确性动机进行操纵能够减少被试鉴别信息时的党派偏见。反之，通过让被试识别其所属群体可能会赞同的信息来对社会性动机进行操纵会降低错误信息鉴别的准确性。这表明准确性动机与社会性

动机具有一定的竞争性。此外，他们的研究还发现，分析性思维与准确性动机对不同变量的影响存在差异。具体来说，准确性动机对党派偏见相关因素的影响权重大于分析性思维，而分析性思维对错误信息鉴别的影响权重大于准确性动机。这不仅揭示了准确性动机与分析性思维的作用差异，也再次验证了社会性动机与准确性动机之间的某种竞争关系。因此，两个理论重点关注的因素并非相互对立，而是属于不同范畴。

综上所述，将系统 2 动机性推理理论与经典推理理论视为相互对立的观点似乎是一种误会。这种误会产生的根本原因在于未能充分区分不同因素在错误信息鉴别中的作用。作为心理学中广泛应用的分析框架，信号检测论能使研究者更清晰地理解不同因素对判断的影响，从而有望化解将两个理论视为对立的误会(Batailler et al., 2022; Gawronski et al., 2023)。

3 信号检测论和相关研究

3.1 信号检测论简介

信号检测论(signal detection theory, SDT)是感知、认知和决策研究中的重要分析框架(Green & Swets, 1966)，其也可以用于理解不同因素如何影响信息鉴别(Gawronski et al., 2023)。通过分析信息鉴别的不同结果(击中、漏报、误报和正确拒绝)，研究者可以深入地理解影响错误信息鉴别的各种因素。

在信号检测论中(Green & Swets, 1966; Lynn & Barrett, 2014)，有两个衡量反应绩效的核心指标：击中率(H)和误报率(FA)。在信息鉴别中，击中率是将真实信息正确地鉴别为真实的次数占真实信息出现总次数的比例。相应的，误报率是将错误信息错误地鉴别为真实的次数占错误信息出现总次数的比例。这两个反应指标为进一步理解影响个体判断的关键因素提供了基础。

信号检测论中区分了影响个体判断的两个关键因素(Green & Swets, 1966; Lynn & Barrett, 2014)：辨别敏感性(d')和判断标准(c)。辨别敏感性反映了个体区分真实信息和错误信息的能力。判断标准反映了个体在判断信息时将信息识别为真或为假的倾向。将击中率和误报率转化为标准分布的 z 分数后，可通过公式$d' = z(H) - z(FA)$计算出辨别敏感性以及公式$c = -0.5 * [z(H) + z(FA)]$计算出判断标准。

在信号检测论中，判断后的收益、事件发生的基本概率和刺激的相似性也是影响个体判

断的重要指标(Lynn & Barrett, 2014)。判断后的收益是指个体在做出特定判断后可能获得的利益或遭受的损失，这种收益或损失的预期会影响个体的判断标准。事件发生的基本概率则涉及到特定事件发生的先验概率，其也会影响个体的判断标准。刺激的相似性是指信号与噪音之间的相似程度，其会影响个体的辨别敏感性。刺激之间的相似性越高，个体区分它们的难度也越大，从而会降低辨别敏感性。相反，刺激之间的相似性越低，个体越容易区分它们，从而会提高辨别敏感性。

3.2　基于信号检测论的错误信息鉴别现有研究

Batailler 等人(2022)指出，先前的研究在探讨人们为何会相信错误信息时，将准确鉴别信息的能力和鉴别信息的反应偏差混为一谈，可能导致了错误结论。因此，他们将信号检测论应用于错误信息鉴别的研究中，以更好地理解个体为何会相信错误信息。Batailler 等人(2022)采用信号检测论的方法重新分析了 Pennycook 等人之前发布的两项研究数据(Pennycook et al., 2018; Pennycook & Rand, 2019)。结果与 Batailler 等人(2022)最初的假设一致，即党派偏见是通过影响判断标准来影响信息鉴别的。具体来说，个体倾向于将与自己党派立场不一致的信息视为错误信息。虽然他们最初假设分析性思维会通过影响个体的辨别敏感性来影响信息鉴别，但结果显示，分析性思维不仅可以通过提高信息辨别敏感性来影响信息鉴别，也可以通过增加将信息鉴别为错误的倾向来影响信息鉴别。此外，他们的研究还进一步探讨了分析性思维与党派偏见之间的关系。结果表明，尽管两者会共同影响错误信息鉴别，但分析性思维的增加并没有减少党派偏见。这进一步支持了本文的观点，即系统 2 动机性推理理论与经典推理理论之间的对立实际上是一种误解。

同样，Gawronski 等人(2023)的研究也揭示了系统 2 动机性推理理论与经典推理理论并非相互对立的观点。通过信号检测论分析发现，在错误信息鉴别时，给予被试更多的分析性思维时间会提高他们的辨别敏感性，同时也会增加将信息鉴别为错误的倾向，但不会减少他们的党派偏见。这一发现与 Batailler 等人(2022)的研究结果一致，再次证明了党派偏见和分析性思维并不是相互影响的。同时，研究还发现党派偏见会通过改变个体的判断标准来影响信息鉴别。具体来说，个体在信息鉴别时会更倾向于信任与其党派立场一致的信息。他们的综合分析发现，党派偏见在错误信息易感性方面发挥着核心作用。虽然未能证实对自我肯定或自我危险的操纵会影响党派偏见，但该研究提示了未来需要进一步探讨影响党派偏见的深

层次因素。

4 基于信号检测论的错误信息鉴别层级模型

上述基于信号检测论的研究共同指出，系统 2 动机性推理理论与经典推理理论并非相互对立(Batailler et al., 2022; Gawronski et al., 2023)。错误地将两者视为对立观点的部分原因在于先前研究未能充分区分不同因素在错误信息鉴别中发挥的作用。Gawronski 等人(2023)基于信号检测论的分析结果，提出了一个理解个体为何会信任错误信息的简要层级模型，旨在更清晰地理解不同因素对错误信息鉴别的影响。在该模型的第一层，将影响错误信息鉴别的因素分为信息辨别敏感性和党派偏见。第二层则深入地探讨了影响第一层因素的心理基础。例如缺乏分析性思维和对真相的关注不足会导致信息辨别敏感性的降低；而个体的主观判断信心则会加剧党派偏见。

虽然 Batailler 等人(2022)和 Gawronski 等人(2023)的工作初步勾勒出了基于信号检测论的错误信息鉴别层级模型，但该模型在一些方面有待完善。首先，该模型在考虑影响错误信息鉴别的因素时，尚未纳入如情绪、信息特征等重要因素。如 Batailler 等人(2022)在其研究中指出，后续研究应该通过信号检测论分析方法来考虑信息来源等其他因素对错误鉴别的影响。其次，该模型对党派偏见影响错误信息鉴别的深层次原因分析尚显不足(Gawronski et al., 2023)。此外，该模型主要基于美国两党制的特殊政治背景，其中党派偏见被视为关键研究对象。然而，错误信息鉴别并不仅限于政治领域，也可能在其他多种社会和文化背景下发生。因此，党派偏见在某种程度上只是影响判断标准的特例。鉴于此，本文将补充和完善基于信号检测论的错误信息鉴别模型，以适用于更广泛的应用场景。

4.1 情绪对错误信息鉴别的影响

情绪是影响判断的重要因素之一(Lerner et al., 2015)，一些研究探讨了情绪对错误信息鉴别的影响(Bago et al., 2022; Martel et al., 2020; Weeks, 2015)。

首先，情绪能够影响个体在鉴别错误信息时的判断标准。Weeks (2015)的研究探讨了情绪状态对错误信息鉴别的影响。该研究通过让被试撰写关于移民改革和死刑的文章来操纵其愤怒或焦虑情绪。随后，被试被要求鉴别含有党派立场的错误信息。结果发现，愤怒情绪增

强了个体支持党派立场信息的倾向，而焦虑情绪则降低了这一倾向。这可能是由于焦虑会促使个体从多角度考虑信息，从而降低了信息鉴别时的党派偏见。相反，愤怒激发了个体的防御心理，减少了对相反信息的思考，从而增加了信息鉴别时的党派偏见。因此，情绪可以通过加强或降低个体信任其所属党派立场信息的倾向性，进而影响判断标准。

其次，情绪也可以作为信息鉴别时的线索和依据，进而影响个体信息鉴别时的辨别敏感性。情感信息理论(feelings-as-information theory)指出(Schwarz & Clore, 2003)，人们会将自身的情绪感受(无论是有意识或无意识的)作为判断的信息来源，对个体的认知风格和判断产生影响。同样，情绪启发式理论强调(Slovic et al., 2007)，个体在面对判断任务时产生的主观情绪体验可以作为判断依据，且这种依赖情感的判断是系统 1 处理的重要成分。在双系统理论视角下，依赖情绪的判断和依赖理性的判断存在一定的对立关系。Martel 等人(2020)在研究中通过诱导提示范式将被试分为情感诱导组、理性归纳组和对照组，旨在探讨错误信息鉴别中依赖情感和依赖理性的不同作用。结果发现，在鉴别信息时，依赖理性的方法比依赖情感更为准确。相关研究也同样表明，在错误信息鉴别时，情绪的参与通常预示着错误信息鉴别准确性的降低(Bago et al., 2022; Martel et al., 2020)。这些研究表明，情绪作为信息鉴别时的线索和依据，可以通过抑制或阻碍个体的理性认知，进而导致错误信息鉴别的辨别敏感性降低。

上述这些证据表明，情绪对错误信息鉴别的影响路径是双重的：一方面，情绪通过影响个体的判断标准，进而影响错误信息鉴别；另一方面，情绪还会通过影响个体的辨别敏感性，进而影响错误信息鉴别。Bago 等人(2022)的补充分析中也验证了这一观点。

4.2 信息特征对错误信息鉴别的影响

错误信息与真实信息存在明显的特征差异，这些特征差异可作为个体错误信息鉴别的线索和依据。根据 Vosoughi 等人(2018)的大数据研究，错误信息通常具有更快、更深、更远、更广的传播特征，内容更新颖，也更容易引发个体的厌恶情绪。刘雅辉等人(2018)的研究显示，错误信息内容中更倾向于包含消极情感词汇、动词、第一人称代词、脏话、非标准语法、符号辅助和加强语气。Li 等人(2022)的研究揭示了错误信息内容的几个关键特征，包括夸张和绝对化的表达方式、诱惑性的措辞、过度强烈的语气、声称信息具有独特性和秘密性、在标点符号使用上存在问题等。此外，该研究还指出，错误信息的排版和呈现方式也具有一定

的不合理性。Zheng 等人(2021)发现，错误信息标题往往倾向于使用确定性的表述，同时也包含更多的情感诉求。吴诗苑等人(2022)从多个角度综述了错误信息的特性，包括内容、用户和传播，指出了错误信息鉴别可分为实体特征识别 (如图形、正文)以及非实体特征识别(如信息的态度、情感和立场)。

信息来源属性特征也可以作为个体错误信息鉴别的线索和依据。根据 Hovland 提出的来源可信度模型，个体对信息发布来源的可信度感知与信息可信度感知呈正相关(Ismagilova et al., 2020)，并进一步将信息来源可信度分为真实性和专业性两个维度。Pilditch 等人(2020)通过概率学习范式的研究，探讨了个体的直接经验和信息来源可信度对信息鉴别的影响。在他们的研究中，信息来源参与任务的次数被用来操纵专业性，将信息来源的利益与被试判断的绩效关联起来作为真实性操纵，同时通过个体选择后的药物有效性反馈作为直接经验操纵。结果发现，个体对信息来源的可信度感知对判断的影响大于直接经验。另外，Trivedi 等人(2020)的研究也验证了信息来源可信度感知与个体信息信任之间的正相关关系。此外，他们的研究通过眼动指标发现(Trivedi et al., 2022)，错误信息鉴别能力较低的个体在信息鉴别时更关注信息来源，这可能是因为他们更倾向于采用信息来源启发式来鉴别信息。这些研究表明，个体倾向于将高可信度信息来源所发布的信息视为真实信息，而将低可信度信息来源所发布的信息视为错误信息。

当个体采用上述信息特征作为错误信息鉴别的线索和依据时，错误信息与真实信息的特征差异大小会影响错误信息鉴别。同样，信号检测论中，刺激的相似性也是影响判断的关键指标，其会直接影响个体的辨别敏感性(Green & Swets , 1966; Lynn & Barrett, 2014)。因此，上述错误信息与真实信息之间的特征差异可以看作是信号检测论框架下信号与噪音之间的物理特征差异。当错误信息与真实信息的特征差异越小，个体的辨别敏感性就越低，区分错误信息与真实信息也更困难。例如，一些错误信息发布者会伪装成高可信度信息来源传播错误信息(Trivedi et al., 2020)，这会降低被试鉴别错误信息的准确性。相反，当错误信息与真实信息的特征差异越大，个体的辨别敏感性就越高，也更易于区分错误信息与真实信息。例如，过于夸张的错误信息标题会提高个体信息鉴别的准确性。

4.3 个体立场以及深层次动机对错误信息鉴别的影响

个体在信息鉴别时所展现的党派偏见，实际上只是基于动机性推理的表面现象，有必要

进一步探讨党派偏见背后的深层次动机。此外，党派偏见只是影响判断标准的一个特例，还应探讨其他类似因素，以更全面地理解错误信息鉴别的影响因素，并有效地拓展和应用现有研究成果。

Lewandowsky 和 Oberauer (2016)认为，错误信息及其负面影响的产生与特定的时代背景密切相关，包括社会资本的减少、价值观的转变、美国社会的分化加剧、科学信任的下降、不同政党对错误信息的易感性差异以及媒体格局的变化等。他们强调，个体对科学信息的拒绝和对错误信息的信任深受其核心信念和世界观的影响。同样，Hornsey 和 Fielding (2017)将个体对信息鉴别的具体动机和态度比喻为树叶和树枝，而将潜在的恐惧、社会认同需求、利益、身份认同、阴谋论、意识形态和世界观等深层次动机比作树根。他们认为，提升错误信息鉴别准确性的关键在于理解并利用这些深层次动机进行有效说服和引导。此外，Van Bavel 和 Pereira (2018)等人提出了一个基于身份信念的模型来解释个体信念为何倾向于党派而非真相。该模型认为，党派群体能够满足个体的归属感、独特性、认知封闭性、权利和资源的获取以及道德价值观的认同。因此，个体可能认为支持真相的收益小于支持党派立场的收益，从而倾向于采用党派立场来鉴别信息。这一观点也与我们上述的假设相吻合，即在错误信息鉴别中，党派立场动机与准确性动机是对立的。这些理论表明，党派立场只是影响个体错误信息鉴别的表面现象，而隐藏在党派立场背后的实际利益与心理需求等深层次动机才是影响个体判断标准的核心要素。在信号检测论视角下(Green & Swets , 1966; Lynn & Barrett, 2014)，这些深层次动机可以看作是判断后的收益，其会通过影响个体的判断标准，进而影响错误信息鉴别。

此外，党派立场作为现有研究焦点，主要是源于美国两党制的特殊政治背景。将党派立场视为模型的核心因素可能会限制模型的广泛适用性。因此，应该将党派偏见作为影响判断标准的一个特例，并考虑其他类似立场（如性别、经济地位、社会角色等）来拓展和应用现有研究成果。上述对党派立场背后深层次动机的探索，也为理解和解释影响判断标准的不同立场提供了理论支持。例如，错误信息中也会涉及性别或经济地位等不同立场，个体在信息鉴别时倾向于信任或拒绝某一立场的信息可能同样源于对实际利益的考量和心理需求。

4.4 基于信号检测论的错误信息鉴别层级模型图

基于信号检测论分析视角，现有研究提供了模型的基础框架(Batailler et al., 2022;

Gawronski et al., 2023)。在这一框架下，错误信息鉴别的影响因素可以通过两个主要途径产生作用：一是影响辨别敏感性，二是影响判断标准。具体而言，缺乏分析性思维和对真相的关注不足会削弱个体在错误信息鉴别时的辨别敏感性；党派偏见则通过改变判断标准来发挥作用，使个体更倾向于将与其立场不一致的信息视为错误信息。

进一步地，本文深入挖掘了党派偏见影响错误信息鉴别的深层次原因。在错误信息鉴别中，党派偏见仅是影响个体判断标准的表面现象和特例，而其背后的实际利益和心理需求等深层次动机才是影响个体判断标准的核心因素。同时，本文还将党派偏见的概念拓展为更广泛的立场，使模型能有效地适用于不同情景和文化背景。

此外，情绪和信息特征等因素也被纳入了该模型。情绪作为影响错误信息鉴别的关键因素，不仅可以通过加强或抑制个体的立场倾向来影响个体的判断标准(Weeks, 2015)，还可以通过与理性思维竞争来降低个体的辨别敏感性(Bago et al., 2022; Slovic et al., 2007)。同时，个体的立场及其背后的深层次动机也会对情绪产生影响(汪新建 等, 2017; Kahan et al., 2010)。

同样，错误信息与真实信息之间的特征差异也会影响个体的辨别敏感性，进而影响错误信息鉴别。当错误信息与真实信息的特征差距越小，也就是错误信息伪装程度较高时，个体的辨别敏感性就越低，也越难准确地鉴别信息。当错误信息与真实信息的特征差距越大，也就是错误信息伪装程度较低时，个体的辨别敏感性就越高，就越能有效地鉴别错误信息。

综上所述，本模型将影响错误信息鉴别的因素分为两大类：一是通过影响判断标准来发挥作用，二是通过影响辨别敏感性来发挥作用。针对判断标准，个体的深层次动机发挥着关键作用，其会使个体更信任与自身立场一致的信息。此外，情绪也会加强或削弱这种立场倾向，进而影响判断标准。同时，个体的立场和背后的深层次动机也能够触发情绪反应。针对辨别敏感性，错误信息与真实信息之间的特征差异以及理性思维发挥着关键作用。情绪也能够通过与理性思维的相互作用来影响辨别敏感性(见图 1)。

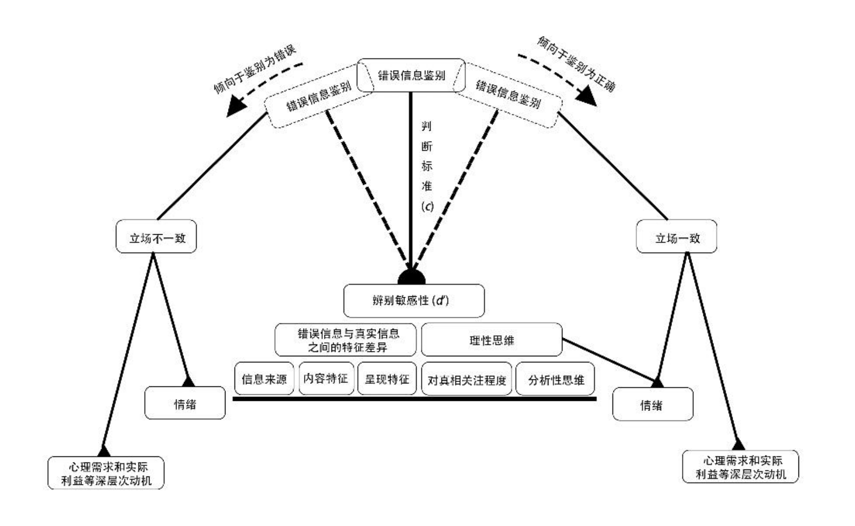


图 1 基于信号检测论的错误信息鉴别层级模型

该模型的建立有助于深入理解不同因素在错误信息鉴别中发挥的作用，并能够调和系统 2 动机性推理理论与经典推理理论之间的分歧。根据基于信号检测论的错误信息鉴别层级模型，系统 2 动机性推理理论与经典推理理论关注的核心因素是通过不同途径来影响错误信息鉴别的。系统 2 动机性推理理论关注的党派偏见主要是通过影响判断标准来产生作用，而经典推理理论关注的分析性思维主要是通过影响辨别敏感性来产生作用。如在 Kahan 等人(2017)的研究结果中发现，具有高计算能力的个体在评估护肤霜效用任务(与党派立场无关)中表现出更高的准确性，但其在禁枪效用任务(与党派立场有关)中更倾向于相信与其党派立场一致的信息，进而降低了判断准确性。根据模型来说，当判断任务不涉及到党派立场时，高计算能力的个体具有较高的辨别敏感性，因此判断信息更为准确。当判断任务涉及到立场时，个体的深层次动机会使判断标准倾向于支持有利自己立场的信息，从而降低了准确性。这进一步说明，两个理论关注的主要因素可以分为通过影响辨别敏感性或通过影响判断标准两条不同途径发挥作用。值得注意的是，尽管这样的划分有助于清晰地理解不同因素在错误信息鉴别时发挥的作用，但这并不意味着某个因素只能通过一条路径发挥作用。实际上，某些因素在其它路径上的作用可能相对较小且不稳定，因而在当前模型中未得到充分体现。如 Batailler 等人(2022)和 Gawronski 等人(2023)的研究发现，分析性思维同样能够通过增加个体将信息鉴别为错误的倾向来影响信息鉴别，但这种影响的效应量较小且不甚稳定。此外，他们的研究还发现，在判断标准上，分析性思维与党派立场的交互作用并不显著(Batailler et al., 2022; Gawronski et al., 2023)，这不仅削弱了分析性思维对判断标准的影响，也一定程度上反驳了分析性思维会加强党派立场的假设。总的来说，系统 2 动机性推理理论与经典推理理论并非彼此对立，而是关注的焦点不同，二者侧重关注的因素均在错误信息鉴别中扮演了重要的角色。

此外，该模型通过区分影响错误信息鉴别的两条途径，不仅有助于理解个体为何会信任错误信息，也为信息治理提供了策略启示。具体而言，针对不同类型的错误信息，可采用不同的治理策略。例如，对于那些因缺乏警觉而易于误信的错误信息，可以通过在信息中添加警告标签的方式(Pennycook et al. ,2020)或加强个体的防范意识等策略来提升辨别敏感性(Swire et al., 2017)，进而降低个体对错误信息的信任。对于含有不同立场且难以直接纠正的错误信息，除了要及时安抚信息所引发的情绪外，还需要深入剖析信息背后隐含的利益关系和心理需求等深层次动机(Hornsey & Fielding, 2017)，并引导个体建立正确的判断标准，进而有效地减少错误信息对个体和社会的影响。

5　总结与展望

本文介绍了影响错误信息鉴别的两个重要理论，即系统 2 动机性推理理论(Kahan, 2013; Kahan & Corbin, 2016; Kahan et al., 2012; Kahan et al., 2017)和经典推理理论(Pennycook & Rand, 2019, 2020)。尽管一些研究将这两个理论视为相互对立的观点(Bago et al., 2020; Gawronski et al., 2023; Pehlivanoglu et al., 2022; Pennycook & Rand, 2019, 2020)，但这种观点可能是一个误会，其原因在于未能充分地理解动机与认知能力在错误信息鉴别中所发挥的不同作用。对此，基于信号检测论的错误信息鉴别层级模型将影响错误信息鉴别的因素分为通过影响判断标准或影响辨别敏感性两条途径产生作用，使研究者更清晰和深入地理解了不同因素在错误信息鉴别中发挥的作用(Batailler et al., 2022; Gawronski et al., 2023)，并调和了现有理论之间的矛盾。此外，本文根据相关研究和理论，深入挖掘了党派偏见的深层次原因，并将党派偏见这一特例拓展为更广泛的立场，以适用于不同情景和文化背景。还有，本文将情绪和信息特征因素也纳入了模型之中，补充和完善了基于信号检测论的错误信息鉴别层级模型，为研究者理解个体如何鉴别错误信息提供了更为全面的理论框架，也为实际应用中的信息鉴别和信息治理提供了理论支持。

尽管基于信号检测论的模型为研究和理解错误信息鉴别提供了新的视角，但其仍存在一些不足。就整体来说，采用基础的信号检测论分析不可避免地存在一些局限性。第一，本模型采用的是基于二元判断("真"或"假")的基础信号检测论，但部分研究采用多点计分来评估对错误信息的信任程度，这会导致模型无法充分利用信息。尽管存在基于多点计分的信号检测论，但这些模型通常较为复杂(如会出现多个判断标准)，需要更深入地探讨相关因素以适应复杂模型。第二，信号检测论的计算过程较为复杂，这可能会导致实际研究中的数据无法较好地拟合模型。第三，信号检测论假设信号和噪音的统计特点是已知稳定的，但现实中错误信息的形态和特性却是在不断变化的。

本文虽探讨了情绪对错误信息鉴别的影响，但仍有许多方面需要进一步研究。首先，尽管本文指出情绪能够通过影响辨别敏感性或判断标准来影响错误信息鉴别，但对于不同情绪的具体作用及其机制尚不清晰。评价倾向框架理论认为(Lerner et al., 2015)，不同情绪会导致个体在信息处理时产生不同的认知模式，但目前的研究尚不充分。例如，积极情绪似乎对人类的推理有着复杂的影响(Pham, 2006)，但探讨积极情绪对错误信息鉴别影响的研究相对匮乏。此外，在现有研究中，愤怒情绪对错误信息鉴别的作用呈现出不一致的结果。Weeks (2015)

的研究发现，愤怒状态会加强个体信任与其党派立场一致信息的倾向。但 Bago 等人(2022)的研究发现，当人们在面对与自己立场不一致的信息标题而感到愤怒时，他们的错误信息鉴别表现反而更准确。这需要未来研究进一步探讨其中的缘由。其次，现有研究探讨情绪与错误信息鉴别关系时，并没有明确区分整合情绪和偶然情绪(朱月龙 等, 2017)。整合情绪是由当前信息引起的主观情绪体验，而偶然情绪是与信息本身相关较小的背景情绪。如 Weeks (2015)和 Bago 等人(2022)的研究中测量的是与错误信息有关的整合情绪。而 Martel 等人(2020)的研究中测量的是与错误信息相关较小的偶然情绪。因此，未来研究可以进一步探讨整合情绪和偶然情绪对错误信息鉴别的影响。

再有，研究可以进一步细化影响错误信息鉴别的认知因素。理性思维是影响错误信息鉴别的重要因素之一，其测量可分为有关认知能力的测量和有关认知倾向的测量(冯雪, 彭凯平, 2015)。同样，在 Stanovich (2016)的三方智力理论中，系统 2 的理性认知可以进一步分为有关计算能力的认知和有关理性思维倾向的认知。正如本文提到的，将系统 2 动机性推理理论和经典推理理论错误地视为对立的另一个原因可能是相关研究中使用的测量工具与其理论所指的对应不够清晰。虽然认知反思测验(Cognitive Reflection Test, CRT)可以测量系统 2 中对直觉反应的抑制能力，但其与个体的算术能力关系密切，且在某些群体中存在天花板效应和地板效应(Thomson & Oppenheimer, 2016)。这提示研究者有必要使用其他测量理性思维的工具来探讨其与错误信息鉴别的关系。已有研究发现，使用积极开放性思维量表作为理性思维倾向的测量指标，不仅能有效地预测错误信息鉴别的准确性(Mirhoseini et al., 2023; Saltor et al., 2023)，还能解释具有妄想倾向的个体为何更易信任错误信息(Bronstein et al., 2019)。值得注意的是，虽然 Kahan (2017)的研究中发现具有高积极开放性思维的个体在鉴别信息时更倾向于信任与自己立场一致的信息(Kahan & Corbin, 2016)，但该结果受到了其他研究者的质疑(Baron, 2017)。此外，Baron (2017)还指出，积极开放性思维量表和认知反思测验所测量的认知风格是不同的，这种差异可能体现在思考的数量和方向上。因此，需要进一步探讨由不同工具所测量的理性思维对错误信息鉴别的影响。

还有，未来研究者可以通过实验来操纵错误信息的特征，从而更清晰地揭示这些特征对错误信息鉴别的影响。目前，错误信息特征主要是通过大数据技术和计算机算法提取的，这种方式虽然具有较高的生态效度，但在因果关系的证明上不够严谨。此外，也有研究者建议通过实验设计对判断后的收益、事件发生的基本概率和刺激的相似性进行操纵，以进一步完善信号检测论(Green & Swets, 1966; Lynn & Barrett, 2014)。因此，后续研究者可以通过实验方法来操纵错误信息与真实信息的特征差异，以进一步完善和验证现有模型。

尽管本模型强调了需要考虑包括性别、经济地位和社会角色在内的多元立场对判断标准的影响，但现有研究多集中于美国两党制政治背景下的党派立场，对于其他立场的研究则相对欠缺。鉴于此，未来研究可以深入探讨不同情境和文化背景下错误信息鉴别中的多元立场，以增强模型的适用性。

总之，基于信号检测论的错误信息鉴别层级模型为研究者提供了一个清晰的理论框架，有助于深入理解和分析影响错误信息鉴别的多种因素，并为实际生活中的信息鉴别和信息治理实践提供了理论依据。

参考文献

冯雪，彭凯平. (2015). 技能和风格:理性思维的两种测量途径. *心理科学进展, 23*(9), 1550–1559.

李艳红，刘佳诺.(2022).人们为什么相信假新闻：对"假新闻信念"的认知心理学解释. *新闻界, 8*, 14–26.

刘雅辉，靳小龙，沈华伟，鲍鹏，程学旗. (2018). 社交媒体中的谣言识别研究综述. *计算机学报, 41*(7), 1536–1558.

彭知辉. (2022). 论中国语境下Disinformation概念的对接、转换与重新阐释. *情报理论与实践, 45*(1), 1–10.

汪新建，张慧娟，武迪，吕小康.(2017).文化对个体风险感知的影响:文化认知理论的解释. *心理科学进展,25*(8),1251–1260.

吴诗苑，董庆兴，宋志君，张斌. (2022). 社交媒体中错误信息的检测方法研究述评. *情报学报, 41*(6), 651–661.

朱月龙，张开华，段锦云. (2017). 建议采纳的情绪机制. *心理科学进展, 25*(9), 1607–1613.

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General, 149*(8), 1608–1613. doi: 10.1037/xge0000729

Bago, B., Rosenzweig, L. R., Berinsky, A. J., & Rand, D. G. (2022). Emotion may predict susceptibility to fake news but emotion regulation does not seem to help. *Cognition and Emotion, 36*(6), 1166–1180. doi: 10.1080/02699931.2022.2090318

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition, 158*, 90–109. doi: 10.1016/j.cognition.2016.10.014

Bago, B., & De Neys, W. (2019). The intuitive greater good: Testing the corrective dual process model of moral cognition. *Journal of Experimental Psychology: General, 148*(10), 1782–1801. doi: 10.1037/xge0000533

Baron, J. (2017). Comment on Kahan and Corbin: Can polarization increase with actively open-minded thinking? *Research & Politics, 4*(1), 1–4. doi: 10.1177/2053168016688122

Batailler, C., Brannon, S. M., Teas, P. E., & Gawronski, B. (2022). A signal detection approach to understanding the identification of fake news. *Perspectives On Psychological Science, 17*(1), 78–98. doi: 10.1177/1745691620986135

Brashier, N. M., & Marsh, E. J. (2020). Judging truth. *Annual Review of Psychology, 71*, 499–515. doi: 10.1146/annurev-psych-010419-050807

Bronstein, M. V., Pennycook, G., Bear, A., Rand, D. G., & Cannon, T. D. (2019). Belief in fake news is associated with delusionality, dogmatism, religious fundamentalism, and reduced analytic thinking. *Journal of Applied*

Research in Memory and Cognition, 8(1), 108–117. doi: 10.1016/j.jarmac.2018.09.005

Evans, J., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives On Psychological Science, 8*(3), 223–241. doi: 10.1177/1745691612460685

Gawronski, B., Ng, N. L., & Luke, D. M. (2023). Truth sensitivity and partisan bias in responses to misinformation. *Journal of Experimental Psychology: General, 152*(8), 2205–2236. doi: 10.1037/xge0001381

Green, D. M., & Swets, J. A. (Eds). (1966). *Signal detection theory and psychophysics*. John Wiley.

Hornsey, M. J., & Fielding, K. S. (2017). Attitude roots and Jiu Jitsu persuasion: Understanding and overcoming the motivated rejection of science. *American Psychologist, 72*(5), 459–473. doi: 10.1037/a0040437

Ismagilova, E., Slade, E., Rana, N. P., & Dwivedi, Y. K. (2020). The effect of characteristics of source credibility on consumer behaviour: A meta-analysis. *Journal of Retailing and Consumer Services, 53*, 101736. doi: 10.1016/j.jretconser.2019.01.005

Kahan, D. M., Braman, D., Monahan, J., Callahan, L., & Peters, E. (2010). Cultural cognition and public policy: The case of outpatient commitment laws. *Law and Human Behavior, 34*(2), 118–140. doi： 10.1007/s10979-008-9174-4

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making, 8*(4), 407–424. doi: 10.1017/S1930297500005271

Kahan, D. M. (2017). 'Ordinary science intelligence': a science-comprehension measure for study of risk and science communication, with notes on evolution and climate change. *Journal of Risk Research, 20*(8), 995–1016. doi: 10.1080/13669877.2016.1148067

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy, 1*(1), 54–86. doi: 10.1017/bpp.2016.2

Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change, 2*(10), 732–735. doi: 10.1038/nclimate1547

Kahan, D. M., & Corbin, J. C. (2016). A note on the perverse effects of actively open-minded thinking on climate-change polarization. *Research & Politics, 3*(4), 1–5. doi:10.1177/2053168016676705

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480–498. doi: 10.1037/0033-2909.108.3.480

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. *Annual Review of Psychology, 66*, 799–823. doi: 10.1146/annurev-psych-010213-115043

Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the "Post-Truth" era. *Journal of Applied Research in Memory and Cognition, 6*(4), 353–369. doi: 10.1016/j.jarmac.2017.07.008

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106–131. doi: 10.1177/1529100612451018

Lewandowsky, S., & Oberauer, K. (2016). Motivated rejection of science. *Current Directions in Psychological Science, 25*(4), 217–222. doi: 10.1177/0963721416654436

Li, Y., Fan, Z., Yuan, X., & Zhang, X. (2022). Recognizing fake information through a developed feature scheme: A user study of health misinformation on social media in China. *Information Processing & Management, 59*(1), Article 102769. doi: 10.1016/j.ipm.2021.102769

Lynn, S. K., & Barrett, L. F. (2014). "Utilizing" signal detection theory. *Psychological Science, 25*(9), 1663–1673. doi: 10.1177/0956797614541991

Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive*

Research: Principles and Implications, 5, Article 47. doi: 10.1186/s41235-020-00252-3

Mirhoseini, M., Early, S., El Shamy, N., & Hassanein, K. (2023). Actively open-minded thinking is key to combating fake news: A multimethod study. *Information & Management, 60*(3), 103761. doi: 10.1016/j.im.2023.103761

Pehlivanoglu, D., Lighthall, N. R., Lin, T., Chi, K. J., Polk, R., Perez, E., Cahill, B. S., & Ebner, N. C. (2022). Aging in an "infodemic": The role of analytical reasoning, affect, and news consumption frequency on news veracity detection. *Journal of Experimental Psychology: Applied, 28*(3), 468–485. doi: 10.1037/xap0000426

Pennycook, G., Bear, A., & Collins, E. (2020). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science, 66*(11), 4944–4957. doi: 10.1287/mnsc.2019.3478

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880. doi: 10.1037/xge0000465

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39–50. doi: 10.1016/j.cognition.2018.06.011

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality, 88*(2), 185–200. doi: 10.1111/jopy.12476

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences, 25*(5), 388–402. doi: 10.1016/j.tics.2021.02.007

Pham, M. (2006). Emotion and rationality: A critical review and interpretation of empirical evidence. *Review of General Psychology, 11*(2), 155–178. doi: 10.1037/1089-2680.11.2.155

Pilditch, T. D., Madsen, J. K., & Custers, R. (2020). False prophets and Cassandra's curse: The role of credibility in belief updating. *Acta Psychologica, 202*, 102956. doi: 10.1016/j.actpsy.2019.102956

Qiu, X. Y., Oliveira, D., Shirazi, A. S., Flammini, A., & Menczer, F. (2017). Limited individual attention and online virality of low-quality information. *Nature Human Behaviour, 1*(8), 1–7. doi: 10.1038/s41562-018-0507-0

Rathje, S., Roozenbeek, J., Van Bavel, J. J., & van der Linden, S. (2023). Accuracy and social motivations shape judgements of (mis)information. *Nature Human Behaviour, 7*(6), 892–903. doi: 10.1038/s41562-023-01540-w

Saltor, J., Barberia, I., & Rodríguez-Ferreiro, J. (2023). Thinking disposition, thinking style, and susceptibility to causal illusion predict fake news discriminability. *Applied Cognitive Psychology, 37*(2), 360–368. doi: 10.1002/acp.4008

Schwarz, N., & Clore, G. L. (2003). Mood as information: 20 years later. *Psychological Inquiry, 14*(3-4), 296–303. doi: 10.1207/S15327965PLI1403&4_20

Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European Journal of Operational Research, 177*(3), 1333–1352. doi: 10.1016/j.ejor.2005.04.006

Stanovich, K. E. (2016). The comprehensive assessment of rational thinking. *Educational Psychologist, 51*(1), 23–34. doi: 10.1080/00461520.2015.1125787

Swire, B., Ecker, U. K., & Lewandowsky, S. (2017). The role of familiarity in correcting inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*(12), 1948–1961. doi: 10.1037/xlm0000422

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning, 20*(2), 215–244. doi: 10.1080/13546783.2013.869763

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making, 11*(1), 99–113. doi: 10.1017/S1930297500007622

Trivedi, N., Krakow, M., Hyatt Hawkins, K., Peterson, E. B., & Chou, W. S. (2020). “Well, the message is from the institute of something”: Exploring source trust of cancer-related messages on simulated facebook posts. *Frontiers in Communication, 5*, 12. doi: 10.3389/fcomm.2020.00012

Trivedi, N., Lowry, M., Gaysynsky, A., & Chou, W. S. (2022). Factors associated with cancer message believability: A mixed methods study on simulated facebook posts. *Journal of Cancer Education, 37*(6), 1870–1878. doi: 10.1007/s13187-021-02054-7

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90*(4), 293–315. doi: 10.1037/0033-295X.90.4.293

Van Bavel, J. J., & Pereira, A. (2018). The partisan brain: An identity-based model of political belief. *Trends in Cognitive Sciences, 22*(3), 213–224. doi: 10.1016/j.tics.2018.01.004

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151. doi: 10.1126/science.aap9559

Wang, Y., McKee, M., Torbica, A., & Stuckler, D. (2019). Systematic literature review on the spread of health-related misinformation on social media. *Social Science & Medicine, 240*, 112552. doi: 10.1016/j.socscimed.2019.112552

Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Strasbourg, France: Council of Europe.

Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication, 65*(4), 699–719. doi: 10.1111/jcom.12164

Zheng, X., Wu, S., & Nie, D. (2021). Online health misinformation and corrective messages in China: A comparison of message features. *Communication Studies, 72*, 1–16. doi: 10.1080/10510974.2021.1917437

Hierarchy model of misinformation identification based on signal detection theory

CAO Chengxu, QI Shisan, JIN Tonglin, ZENG Xiaoye, AN Yeqing, BU Tana

(School of Psychology, Inner Mongolia Normal University, Key research base of psychological education in Inner Mongolia Colleges and Universities, Hohhot 010022, China)

Abstract: In the field of misinformation identification research, the motivated System 2 reasoning and classical reasoning accounts probe the influencing factors that shape individuals' ability to identify misinformation from different perspectives, yet diverge in their interpretations of cognitive abilities' roles. Building upon existing research, this study introduces factors such as emotions, information characteristics, individual stances, and their underlying motivations to further refine the hierarchical model of misinformation identification based on a signal detection theory. The objective is to enrich our comprehension of the multifaceted ways in which these diverse elements bear upon the process of misinformation identification. By differentiating the

influence of various factors on both the discrimination sensitivity and the judgment criteria within the identification process, the model not only reconciles the contrasting perspectives on cognitive abilities posited by motivated System 2 reasoning and classical reasoning accounts but also furnishes a detailed and systematically organized analytical framework. This framework is instrumental in elucidating the intricate mechanisms that underpin the identification of misinformation.

Key words: misinformation, signal detection theory, affecting factors, motivated System 2 reasoning account, classical reasoning account